(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 0 854 188 A2

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
22.07.1998 Bulletin 1998/30

(21) Application number: **98300361.7**

(22) Date of filing: **20.01.1998**

(51) Int Cl.[6]: **C12N 15/53**, C12Q 1/68, C12N 15/62, C07K 16/12, G01N 33/566, A61K 38/44

(84) Designated Contracting States:
AT BE CH DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE

(30) Priority: **21.01.1997 US 36125 P**

(71) Applicant: **Smithkline Beecham**
Philadelphia, PA 19103 (US)

(72) Inventor: **Burnham, Martin K.R.,**
SmithKline Beecham Pharm.
Collegeville, PA 19426-0989 (US)

(74) Representative: **Mallalieu, Catherine Louise**
D. Young & Co.,
21 New Fetter Lane
London EC4A 1DA (GB)

(54) Streptococcus pneumoniae aroE polypeptides and polynucleotides

(57) The invention provides aroE polypeptides and DNA (RNA) encoding aroE polypetides and methods for producing such polypeptides by recombinant techniques. Also provided are methods for utilizing aroE polypeptide for the protection against infection. particularly bacterial infections.

EP 0 854 188 A2

Printed by Jouve, 75001 PARIS (FR)

In accordance with certain preferred embodiments of this aspect of the invention, there are provided products, compositions and methods for (i) assessing aroE expression, (ii) treating disease for example, otitis media, conjunctivitis, pneumonia, bacteremia, meningitis, sinusitis, pleural empyema and endocarditis, and most particularly meningitis, such as for example infection of cerebrospinal fluid., (iii) assaying genetic variation, (iv) and administering a aroE polypeptide or polynucleotide to an organism to raise an immunological response against a bacteria, especially a *Streptococcus pneumoniae* bacteria.

In accordance with certain preferred embodiments of this and other aspects of the invention there are provided polynucleotides that hybridize to aroE polynucleotide sequences, particularly under stringent conditions.

In certain preferred embodiments of this aspect of the invention there are provided antibodies against aroE polypeptides

In accordance with another aspect of the invention, there are provided aroE agonists and antagonists each of which are also preferably bacteriostatic or bacteriocidal.

In a further aspect of the invention there are provided compositions comprising a aroE polynucleotide or a aroE polypeptide for administration to a cell or to a multicellular organism.

Various changes and modifications within the spirit and scope of the disclosed invention will become readily apparent to those skilled in the art from reading the following descriptions and from reading the other parts of the present disclosure.

BRIEF DESCRIPTION OF THE DRAWINGS

The following drawings depict certain embodiments of the invention. They are illustrative only and do not limit the invention otherwise disclosed herein.

Figure 1 shows the polynucleotide sequence of *Streptococcus pneumoniae* aroE [SEQ ID NO:1].

Figure 2 shows the amino acid sequence of *Streptococcus pneumoniae* aroE [SEQ ID NO:2] deduced from the polynucleotide sequence of Figure 1.

GLOSSARY

The following definitions are provided to facilitate understanding of certain terms used frequently herein.

"Host cell" is a cell which has been transformed or transfected, or is capable of transformation or transfection by an exogenous polynucleotide sequence.

"Identity," as known in the art, is a relationship between two or more polypeptide sequences or two or more polynucleotide sequences, as determined by comparing the sequences. In the art, "identity" also means the degree of sequence relatedness between polypeptide or polynucleotide sequences, as the case may be, as determined by the match between strings of such sequences. "Identity" and "similarity" can be readily calculated by known methods *(Computational Molecular Biology,* Lesk, A.M., ed., Oxford University Press, New York, 1988; *Biocomputing: Informatics and Genome Projects,* Smith, D.W., ed., Academic Press, New York, 1993; *Computer Analysis of Sequence Data,* Part 1, Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, 1994; *Sequence Analysis in Molecular Biology,* von Heinje, G., Academic Press, 1987; and *Sequence Analysis Primer,* Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991). While there exist a number of methods to measure identity and similarity between two sequences, both terms are well known to skilled artisans *(Sequence Analysis in Molecular Biology,* von Heinje, G., Academic Press, 1987; *Sequence Analysis Primer,* Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991; and Carillo, H., and Lipman, D., SIAM *J. Applied Math., 48: 1073* (1988). Methods commonly employed to determine identity or similarity between sequences include, but are not limited to those disclosed in Carillo, H., and Lipman, D., SIAM *J Applied Math., 48:1073* (1988). Preferred methods to determine identity are designed to give the largest match between the sequences tested. Methods to determine identity and similarity are codified in publicly available computer programs. Preferred computer program methods to determine identity and similarity between two sequences include, but are not limited to, GCG program package (Devereux, J., et al., *Nucleic Acids Research 12(1).* 387 (1984)), BLASTP, BLASTN, and FASTA (Atschul, S.F. et al., *J. Molec. Biol. 215.* 403-410 (1990)). The BLAST X program is publicly available from NCBI and other sources *(BLAST Manual,* Altschul, S., *et al.,* NCBI NLM NIH Bethesda, MD 20894, Altschul, S., *et al., J. Mol. Biol. 215:* 403-410 (1990)).

"Isolated" means altered "by the hand of man" from its natural state, *i.e.*, if it occurs in nature, it has been changed or removed from its original environment, or both. For example, polynucleotide or a polypeptide naturally present in a living organism is not "isolated," but the same polynucleotide or polypeptide separated from the coexisting materials of its natural state is "isolated", as the term is employed herein.

"Polynucleotide(s)" generally refers to any polyribonucleotide or polydeoxribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. "Polynucleotide(s)" include, without limitation, single-and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions or single-, double- and triple-stranded regions, single-

which is related by amino acid sequence homology to known *Escherichia coli* Shikimate 5-Dehydrogenase polypeptide. The invention relates especially to aroE having the nucleotide and amino acid sequences set out in Figure 1 and Figure 2 respectively, and to the aroE nucleotide sequences of the DNA in NCIMB Deposit No. 40794 and amino acid sequences encoded thereby.

Techniques are available to evaluate temporal gene expression in bacteria, particularly as it applies to viability under laboratory and host infection conditions. A number of methods can be used to identify genes which are essential to survival *per se*, or essential to the establishment and/or maintenance of an infection. Identification of expression of a sequence by one of these methods yields additional information about its function and assists in the selection of such sequence for further development as a screening target. Briefly, these approaches include for example:

1) Signature Tagged Mutagenesis (STM)

This technique is described by Hensel *et al., Science* 269: 400-403(1995), the contents of which is incorporated by reference for background purposes. Signature tagged mutagenesis identifies genes necessary for the establishment/maintenance of infection in a given infection model.

The basis of the technique is the random mutagenesis of target organism by various means (e.g., transposons) such that unique DNA sequence tags are inserted in close proximity to the site of mutation. The tags from a mixed population of bacterial mutants and bacteria recovered from an infected hosts are detected by amplification, radiolabeling and hybridization analysis. Mutants attenuated in virulence are revealed by absence of the tag from the pool of bacteria recovered from infected hosts.

In *Streptococcus pneumoniae*, because the transposon system is less well developed, a more efficient way of creating the tagged mutants is to use the insertion-duplication mutagenesis technique as described by Morrison *et al. J. Bacteriol.* 159:870 (1984) the contents of which is incorporated by reference for background purposes.

2) In Vivo Expression Technology (IVET)

This technique is described by Camilli *et al., Proc. Nat'l. Acad. Sci. USA.* 91:2634-2638 (1994) and Mahan *et al., Infectious Agents and Diseases* 2:263-268 (1994), the contents of each of which is incorporated by reference for background purposes. IVET identifies genes up-regulated during infection when compared to laboratory cultivation, implying an important role in infection. Sequences identified by this technique are implied to have a significant role in infection establishment/maintenance.

In this technique random chromosomal fragments of target organism are cloned upstream of promoter-less reporter gene in a plasmid vector. The pool is introduced into a host and at various times after infection bacteria may be recovered and assessed for the presence of reporter gene expression. The chromosomal fragment carried upstream of an expressed reporter gene should carry a promoter or portion of a gene normally upregulated during infection. Sequencing upstream of the reporter gene allows identification of the up regulated gene.

3) Differential display

This technique is described by Chuang *et al., J. Bacteriol.* 175:2026-2036 (1993), the contents of which is incorporated by reference for background purposes. This method identifies those genes which are expressed in an organism by identifying mRNA present using randomly-primed RT-PCR. By comparing pre-infection and post infection profiles, genes up and down regulated during infection can be identified and the RT-PCR product sequenced and matched to library sequences.

4) Generation of conditional lethal mutants by transposon mutagenesis.

This technique, described by de Lorenzo, V. *et al., Gene* 123:17-24 (1993); Neuwald, A. F. *et al., Gene* 125: 69-73 (1993); and Takiff, H. E. *et al., J. Bacteriol.* 174:1544- 1553(1992), the contents of which is incorporated by reference for background purposes, identifies genes whose expression are essential for cell viability.

In this technique transposons carrying controllable promoters, which provide transcription outward from the transposon in one or both directions, are generated. Random insertion of these transposons into target organisms and subsequent isolation of insertion mutants in the presence of inducer of promoter activity ensures that insertions which separate promoter from coding region of gene whose expression is essential for cell viability will be recovered. Subsequent replica plating in the absence of inducer identifies such insertions since they fail to survive. Sequencing of the flanking regions of the transposon allows identification of site of insertion and identification of the gene disrupted. Close monitoring of the changes in cellular processes/morphology during growth in the absence of inducer yields information on likely function of the gene. Such monitoring could include flow cytometry (cell division, lysis, redox

Polypeptides

The polypeptides of the invention include the polypeptide of Figure 2 [SEQ ID NO:2] (in particular the mature polypeptide) as well as polypeptides and fragments, particularly those which have the biological activity of aroE and also those which have at least 70% identity to the polypeptide of Figure 2 [SEQ ID NO:2] or the relevant portion, preferably at least 80% identity to the polypeptide of Figure 2 [SEQ ID NO:2], and more preferably at least 90% similarity (more preferably at least 90% identity) to the polypeptide of Figure 2 [SEQ ID NO:2] and still more preferably at least 95% similarity (still more preferably at least 95% identity) to the polypeptide of Figure 2 [SEQ ID NO:2] and also include portions of such polypeptides with such portion of the polypeptide generally containing at least 30 amino acids and more preferably at least 50 amino acids.

Variants that are fragments of the polypeptides of the invention may be employed for producing the corresponding full-length polypeptide by peptide synthesis; therefore, these variants may be employed as intermediates for producing the full-length polypeptides. Variants that are fragments of the polynucleotides of the invention may be used to synthesize full-length polynucleotides of the invention.

A fragment is a variant polypeptide having an amino acid sequence that entirely is the same as part but not all of the amino acid sequence of the aforementioned polypeptides. As with aroE polypeptides fragments may be "free-standing," or comprised within a larger polypeptide of which they form a part or region, most preferably as a single continuous region, a single larger polypeptide.

Preferred fragments include, for example, truncation polypeptides having a portion of the amino acid sequence of Figure 2 [SEQ ID NO:2], or of variants thereof, except for deletion of a continuous series of residues that includes the amino terminus, or a continuous series of residues that includes the carboxyl terminus or deletion of two continuous series of residues, one including the amino terminus and one including the carboxyl terminus. Degradation forms of the polypeptides of the invention in a host cell, particularly a *Streptococcus pneumoniae,* are also preferred. Also preferred are fragments characterized by structural or functional attributes such as fragments that comprise alpha-helix and alpha-helix forming regions, beta-sheet and beta-sheet-forming regions, turn and turn-forming regions, coil and coil-forming regions, hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta amphipathic regions, flexible regions, surface-forming regions, substrate binding region, and high antigenic index regions.

Also preferred are biologically active fragments which are those fragments that mediate activities of aroE, including those with a similar activity or an improved activity, or with a decreased undesirable activity. Also included are those fragments that are antigenic or immunogenic in an animal, especially in a human.

Polynucleotides

Another aspect of the invention relates to isolated polynucleotides which encode the aroE polypeptide having the deduced amino acid sequence of Figure 2 [SEQ ID NO:2] and polynucleotides closely related thereto and variants thereto.

Using the information provided herein, such as the polynucleotide sequence set out in

Figure 1 [SEQ ID NO:1], a polynucleotide of the invention encoding aroE polypeptide may be obtained using standard cloning and screening, such as those for cloning and sequencing chromosomal DNA fragments from *Streptococcus pneumoniae* 0100993 cells as starting material, followed by obtaining a full length clone. For example, to obtain a polynucleotide sequence of the invention, such as that sequence given in Figure 1 [SEQ ID NO:1], typically a library of clones of chromosomal DNA of *Streptococcus pneumoniae* 0100993 in *E.coli* or some other suitable host is probed with a radiolabeled oligonucleotide, preferably a 17-mer or longer, derived from a partial sequence. Clones carrying DNA identical to that of the probe can then be distinguished using stringent conditions. By sequencing the individual clones thus identified with sequencing primers designed from the original sequence it is then possible to extend the sequence in both directions to determine the full gene sequence. Conveniently such sequencing is performed using denatured double stranded DNA prepared from a plasmid clone. Suitable techniques are described by Maniatis, T., Fritsch, E.F. and Sambrook et al., *MOLECULAR CLONING, A LABORATORY MANUAL,* 2nd Ed.; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1989). (see Screening By Hybridization 1.90 and Sequencing Denatured Double-Stranded DNA Templates 13.70). Illustrative of the invention, the polynucleotide set out in Figure 1 [SEQ ID NO:1] was discovered in a DNA library derived from *Streptococcus pneumoniae* 0100993.

The DNA sequence thus obtained is set out in Figure 1 [SEQ ID NO: 1]. It contains an open reading frame encoding a protein having about the number of amino acid residues set forth in Figure 2 [SEQ ID NO:2] with a deduced molecular weight that can be calculated using amino acid residue molecular weight values well known in the art. aroE of the invention is structurally related to other proteins of the aromatic amino acid biosynthesis family, as shown by the results of sequencing the DNA encoding aroE of the deposited strain. The protein exhibits greatest homology to *Escherichia*

have a high sequence similarity to the aroE gene. Such probes generally will comprise at least 15 bases. Preferably, such probes will have at least 30 bases and may have at least 50 bases. Particularly preferred probes will have at least 30 bases and will have 50 bases or less.

For example, the coding region of the aroE gene may be isolated by screening using the known DNA sequence to synthesize an oligonucleotide probe. A labeled oligonucleotide having a sequence complementary to that of a gene of the invention is then used to screen a library of cDNA, genomic DNA or mRNA to determine which members of the library the probe hybridizes to.

The polynucleotides and polypeptides of the invention may be employed as research reagents and materials for discovery of treatments of and diagnostics for disease, particularly human disease, as further discussed herein relating to polynucleotide assays, *inter alia*.

Polynucleotides ofthe invention that are oligonucleotides derived from the sequences of SEQ ID NOS: 1 and 2 may be used in the processes herein as described, but preferably for PCR, to determine whether or not the polynucleotides identified herein in whole or in part are transcribed in infected tissue. It is recognized that such sequences will also have utility in diagnosis of the stage of infection and type of infection the pathogen has attained.

The polynucleotides may encode a polypeptide which is the mature protein plus additional amino or carboxyl-terminal amino acids, or amino acids interior to the mature polypeptide (when the mature form has more than one polypeptide chain, for instance). Such sequences may play a role in processing of a protein from precursor to a mature form, may allow protein transport, may lengthen or shorten protein half-life or may facilitate manipulation of a protein for assay or production, among other things. As generally is the case in *vivo*, the additional amino acids may be processed away from the mature protein by cellular enzymes.

A precursor protein, having the mature form of the polypeptide fused to one or more prosequences may be an inactive form of the polypeptide. When prosequences are removed such inactive precursors generally are activated. Some or all of the prosequences may be removed before activation. Generally, such precursors are called proproteins.

In sum, a polynucleotide of the invention may encode a mature protein, a mature protein plus a leader sequence (which may be referred to as a preprotein), a precursor of a mature protein having one or more prosequences which are not the leader sequences of a preprotein, or a preproprotein, which is a precursor to a proprotein, having a leader sequence and one or more prosequences, which generally are removed during processing steps that produce active and mature forms of the polypeptide.

Vectors, host cells, expression

The invention also relates to vectors which comprise a polynucleotide or polynucleotides of the invention, host cells which are genetically engineered with vectors of the invention and the production of polypeptides of the invention by recombinant techniques. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the invention.

For recombinant production, host cells can be genetically engineered to incorporate expression systems or portions thereof or polynucleotides of the invention. Introduction of a polynucleotide into the host cell can be effected by methods described in many standard laboratory manuals, such as Davis et al., *BASIC METHODS IN MOLECULAR BIOLOGY*, (1986) and Sambrook et al., *MOLECULAR CLONING: A LABORATORY MANUAL*, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989), such as, calcium phosphate transfection, DEAE-dextran mediated transfection, transvection, microinjection, cationic lipid-mediated transfection, electroporation, transduction, scrape loading, ballistic introduction and infection.

Representative examples of appropriate hosts include bacterial cells, such as streptococci, staphylococci, *E. coli*, streptomyces and *Bacillus subtilis* cells; fungal cells, such as yeast cells and *Aspergillus* cells; insect cells such as *Drosophila* S2 and *Spodoptera* Sf9 cells; animal cells such as CHO, COS, HeLa, C127, 3T3, BHK, 293 and Bowes melanoma cells; and plant cells.

A great variety of expression systems can be used to produce a polypeptide of the invention. Such vectors include, among others, chromosomal, episomal and virus-derived vectors, *e.g.*, vectors derived from bacterial plasmids, from bacteriophage, from transposons, from yeast episomes, from insertion elements, from yeast chromosomal elements, from viruses such as baculoviruses, papova viruses, such as SV40, vaccinia viruses, adenoviruses, fowl pox viruses, pseudorabies viruses and retroviruses, and vectors derived from combinations thereof, such as those derived from plasmid and bacteriophage genetic elements, such as cosmids and phagemids. The expression system constructs may contain control regions that regulate as well as engender expression. Generally, any system or vector suitable to maintain, propagate or express polynucleotides and/or to express a polypeptide in a host may be used for expression in this regard. The appropriate DNA sequence may be inserted into the expression system by any of a variety of well-known and routine techniques, such as, for example, those set forth in Sambrook *et al.*, *MOLECULAR CLONING, A LABORATORY MANUAL, (supra)*.

For secretion of the translated protein into the lumen of the endoplasmic reticulum, into the periplasmic space or

such as. for example, amplification, PCR, RT-PCR, RNase protection. Northern blotting and other hybridization methods

In addition, a diagnostic assay in accordance with the invention for detecting over-expression of aroE protein compared to normal control tissue samples may be used to detect the presence of an infection, for example. Assay techniques that can be used to determine levels of a aroE protein, in a sample derived from a host are well-known to those of skill in the art. Such assay methods include radioimmunoassays, competitive-binding assays, Western Blot analysis and ELISA assays.

Antibodies

The polypeptides of the invention or variants thereof or cells expressing them can be used as an immunogen to produce antibodies immunospecific for such polypeptides. "Antibodies" as used herein includes monoclonal and polyclonal antibodies, chimeric, single chain, simianized antibodies and humanized antibodies, as well as Fab fragments, including the products of an Fab immunolglobulin expression library.

Antibodies generated against the polypeptides of the invention can be obtained by administering the polypeptides or epitope-bearing fragments, analogues or cells to an animal, preferably a nonhuman, using routine protocols. For preparation of monoclonal antibodies, any technique known in the art which provides antibodies produced by continuous cell line cultures can be used. Examples include various techniques, such as those in Kohler, G. and Milstein, C., *Nature 256*: 495-497 (1975); Kozbor *et al., Immunology Today 4*: 72 (1983); Cole et al., pg. 77-96 in *MONOCLONAL ANTIBODIES AND CANCER THERAPY*. Alan R. Liss, Inc. (1985).

Techniques for the production of single chain antibodies (U.S. Patent No. 4,946,778) can be adapted to produce single chain antibodies to polypeptides of this invention. Also, transgenic mice, or other organisms such as other mammals, may be used to express humanized antibodies.

Alternatively phage display technology could be utilized to select antibody genes with binding activities towards the polypeptide either from repertoires of PCR amplified v-genes of lymphocytes from humans screened for possessing anti-Fbp or from naive libraries (McCafferty, J. et al., (1990), Nature **348**, 552-554; Marks, J. et al., (1992) Biotechnology *10*, 779-783). The affinity of these antibodies can also be improved by chain shuffling (Clackson, T. et al., (1991) Nature **352**, 624-628).

If two antigen binding domains are present each domain may be directed against a different epitope - termed 'bispecific' antibodies.

The above-described antibodies may be employed to isolate or to identify clones expressing the polypeptides to purify the polypeptides by affinity chromatography.

Thus, among others, antibodies against aroE may be employed to treat infections, particularly bacterial infections and especially otitis media, conjunctivitis, pneumonia, bacteremia, meningitis, sinusitis, pleural empyema and endocarditis, and most particularly meningitis, such as for example infection of cerebrospinal fluid,.

Polypeptide variants include antigenically, epitopically or immunologically equivalent variants which form a particular aspect of this invention. The term "antigenically equivalent derivative" as used herein encompasses a polypeptide or its equivalent which will be specifically recognised by certain antibodies which, when raised to the protein or polypeptide according to the invention, interfere with the immediate physical interaction between pathogen and mammalian host. The term "immunologically equivalent derivative" as used herein encompasses a peptide or its equivalent which when used in a suitable formulation to raise antibodies in a vertebrate, the antibodies act to interfere with the immediate physical interaction between pathogen and mammalian host.

The polypeptide, such as an antigenically or immunologically equivalent derivative or a fusion protein thereof is used as an antigen to immunize a mouse or other animal such as a rat or chicken. The fusion protein may provide stability to the polypeptide. The antigen may be associated, for example by conjugation, with an immunogenic carrier protein for example bovine serum albumin (BSA) or keyhole limpet haemocyanin (KLH). Alternatively a multiple antigenic peptide comprising multiple copies of the protein or polypeptide, or an antigenically or immunologically equivalent polypeptide thereof may be sufficiently antigenic to improve immunogenicity so as to obviate the use of a carrier.

Preferably the antibody or variant thereof is modified to make it less immunogenic in the individual. For example, if the individual is human the antibody may most preferably be "humanized"; where the complimentarity determining region(s) of the hybridoma-derived antibody has been transplanted into a human monoclonal antibody, for example as described in Jones, P. et al. (1986), Nature **321**, 522-525 or Tempest et al.,(1991) Biotechnology **9**,266,273.

The use of a polynucleotide of the invention in genetic immunization will preferably employ a suitable delivery method such as direct injection of plasmid DNA into muscles (Wolff et al., Hum Mol Genet 1992, 1:363, Manthorpe et al., Hum. Gene Ther. 1963:4, 419), delivery of DNA complexed with specific protein carriers (Wu et al., J Biol Chem 1989:264.16985), coprecipitation of DNA with calcium phosphate (Benvenisty & Reshef, PNAS, 1986:83.9551), encapsulation of DNA in various forms of liposomes (Kaneda et al., Science 1989:243,375), particle bombardment (Tang et al. Nature 1992, 356 152, Eisenbraun et al., DNA Cell Biol 1993, 12:791) and *in vivo* infection using cloned retroviral

produce antibody to protect said individual from infection, particularly bacterial infection and most particularly *Streptococcus pneumoniae* infections. Yet another aspect of the invention relates to a method of inducing immunological response in an individual which comprises, through gene therapy, delivering gene encoding aroE, or a fragment or a variant thereof, for expressing aroE, or a fragment or a variant thereof *in vivo* in order to induce an immunological response to produce antibody to protect said individual from disease.

A further aspect of the invention relates to an immunological composition which, when introduced into a host capable or having induced within it an immunological response, induces an immunological response in such host to a aroE or protein coded therefrom, wherein the composition comprises a recombinant aroE or protein coded therefrom comprising DNA which codes for and expresses an antigen of said aroE or protein coded therefrom.

The aroE or a fragment thereof may be fused with co-protein which may not by itself produce antibodies, but is capable of stabilizing the first protein and producing a fused protein which will have immunogenic and protective properties. Thus fused recombinant protein, preferably further comprises an antigenic co-protein, such as Glutathione-S-transferase (GST) or beta-galactosidase, relatively large co-proteins which solubilise the protein and facilitate production and purification thereof. Moreover, the co-protein may act as an adjuvant in the sense of providing a generalized stimulation of the immune system. The co-protein may be attached to either the amino or carboxy terminus of the first protein.

Provided by this invention are compositions, particularly vaccine compositions, and methods comprising the polypeptides or polynucleotides of the invention and immunostimulatory DNA sequences, such as those described in Sato, Y. *et al.* Science 273: 352 (1996).

Also, provided by this invention are methods using the described polynucleotide or particular fragments thereof which have been shown to encode non-variable regions of bacterial cell surface proteins in DNA constructs used in such genetic immunization experiments in animal models of infection with *Streptococcus pneumoniae* will be particularly useful for identifying protein epitopes able to provoke a prophylactic or therapeutic immune response. It is believed that this approach will allow for the subsequent preparation of monoclonal antibodies of particular value from the requisite organ of the animal successfully resisting or clearing infection for the development of prophylactic agents or therapeutic treatments of bacterial infection, particularly *Streptococcus pneumoniae* infections, in mammals, particularly humans.

The polypeptide may be used as an antigen for vaccination of a host to produce specific antibodies which protect against invasion of bacteria, for example by blocking adherence of bacteria to damaged tissue. Examples of tissue damage include wounds in skin or connective tissue caused e.g. by mechanical, chemical or thermal damage or by implantation of indwelling devices, or wounds in the mucous membranes, such as the mouth, mammary glands, urethra or vagina.

The invention also includes a vaccine formulation which comprises the immunogenic recombinant protein together with a suitable carrier. Since the protein may be broken down in the stomach, it is preferably administered parenterally, including, for example, administration that is subcutaneous, intramuscular, intravenous, or intradermal. Formulations suitable for parenteral administration include aqueous and non-aqueous sterile injection solutions which may contain anti-oxidants, buffers, bacteriostats and solutes which render the formulation instonic with the bodily fluid, preferably the blood, of the individual; and aqueous and non-aqueous sterile suspensions which may include suspending agents or thickening agents. The formulations may be presented in unit-dose or multi-dose containers, for example, sealed ampoules and vials and may be stored in a freeze-dried condition requiring only the addition of the sterile liquid carrier immediately prior to use. The vaccine formulation may also include adjuvant systems for enhancing the immunogenicity of the formulation, such as oil-in water systems and other systems known in the art. The dosage will depend on the specific activity of the vaccine and can be readily determined by routine experimentation.

While the invention has been described with reference to certain aroE, it is to be understood that this covers fragments of the naturally occurring protein and similar proteins with additions, deletions or substitutions which do not substantially affect the immunogenic properties of the recombinant protein.

Compositions, kits and administration

The invention also relates to compositions comprising the polynucleotide or the polypeptides discussed above or the agonists or antagonists. The polypeptides of the invention may be employed in combination with a non-sterile or sterile carrier or carriers for use with cells, tissues or organisms, such as a pharmaceutical carrier suitable for administration to a subject. Such compositions comprise, for instance, a media additive or a therapeutically effective amount of a polypeptide of the invention and a pharmaceutically acceptable carrier or excipient. Such carriers may include, but are not limited to, saline, buffered saline, dextrose, water, glycerol, ethanol and combinations thereof. The formulation should suit the mode of administration. The invention further relates to diagnostic and pharmaceutical packs and kits comprising one or more containers filled with one or more of the ingredients of the aforementioned compositions of the invention.

size-fractionated by either of two methods.

Method 1

Total cellular DNA is mechanically sheared by passage through a needle in order to size-fractionate according to standard procedures. DNA fragments of up to I lkbp in size are rendered blunt by treatment with exonuclease and DNA polymerase, and EcoRI linkers added. Fragments are ligated into the vector Lambda ZapII that has been cut with EcoRI. the library packaged by standard procedures and *E.coli* infected with the packaged library. The library is amplified by standard procedures.

Method 2

Total cellular DNA is partially hydrolyzed with a one or a combination of restriction enzymes appropriate to generate a series of fragments for cloning into library vectors (e.g., RsaI, PalI, AluI, Bsh1235I), and such fragments are size-fractionated according to standard procedures. EcoRI linkers are ligated to the DNA and the fragments then ligated into the vector Lambda ZapII that have been cut with EcoRI, the library packaged by standard procedures. and *E.coli* infected with the packaged library. The library is amplified by standard procedures.

(viii) ATTORNEY/AGENT INFORMATION:

940 (A) NAME: MALLALIEU, Catherine Louise

(B) GENERAL AUTHORISATION NUMBER: 37128

(C) REFERENCE/DOCKET NUMBER: GM50000

(ix) TELECOMMUNICATION INFORMATION:

945 (A) TLEEPHONE: +44 171 353 4343

(B) TELEFAX: -44 171 353 7777

(2) INFORMATION FOR SEQ ID NO:1:

950

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 772 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

955 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

960 ATGAAGCTTG ATGGCTATAC ACGTTTAGCT GCAGTTGTTG CCAATCCTAT TAAGCATTCT 60

ATTTCTCCCT TCATCCACAA TAGAGCCTTT GAGGCGACAG CTACCAACGG TGCTTATGTG 120

GCTTGGGAGA TTGAAGCGAG TGACTTGGTA GAAACAGTGG CCAATATTCG TCGCTACCAG 180

965 ATGTTTGGCA TCAATCTGTC CATGCCCTAT AAGGAGCAGG TGATTCCTTA TTTGGATAAG 240

CTGAGCGATG AAGCGCGCTT GATTGGTGCG GTTAATACGG TTGTCAATGA GAATGGCAAT 300

35 40 45

Leu Val Glu Thr Val Ala Asn Ile Arg Arg Tyr Gln Met Phe Gly Ile

1005 50 55 60

Asn Leu Ser Met Pro Tyr Lys Glu Gln Val Ile Pro Tyr Leu Asp Lys

65 70 75 80

Leu Ser Asp Glu Ala Arg Leu Ile Gly Ala Val Asn Thr Val Val Asn

85 90 95

1010 Glu Asn Gly Asn Leu Ile Gly Tyr Asn Thr Asp Gly Lys Gly Phe Phe

100 105 110

Lys Cys Leu Pro Ser Phe Thr Ile Ser Gly Lys Lys Met Thr Leu Leu

115 120 125

Gly Ala Gly Gly Ala Ala Lys Ser Ile Leu Ala Gln Ala Ile Leu Asp

1015 130 135 140

Gly Val Ser Gln Ile Ser Val Phe Val Arg Ser Val Ser Met Glu Lys

145 150 155 160

Thr Arg Pro Tyr Leu Asp Lys Leu Gln Glu Gln Thr Gly Phe Lys Val

165 170 175

1020 Asp Leu Cys Ala Leu Glu Tyr Val Ser Glu Leu Gln Ala Arg Ile Ala

180 185 190

Glu Ser Asp Leu Leu Val Asn Ala Thr Ser Val Gly Met Asp Gly Gln

195 200 205

Phe Ser Pro Val Pro Glu Asn Ile Val Leu Pro Glu Thr Leu Leu Val

1025 210 215 220

Ala Asp Ile Ile Tyr Gln Pro Phe Glu Thr Pro Phe Leu Lys Trp Ala

225 230 235 240

Arg Ser Gln Gly Asn Pro Ala Val Asn Gly Leu Gly Met Leu Leu Tyr

245 250 255

1030 Gln

(2) INFORMATION FOR SEQ ID NO:3:

1035 (i) SEQUENCE CHARACTERISTICS:

6. The polynucleotide of Claim 2 which encodes a polypeptide comprising amino acid 1 to 257 of SEQ ID NO:2.

7. An isolated polynucleotide comprising a member selected from the group consisting of:

(a) a polynucleotide having at least a 70% identity to a polynucleotide encoding the same mature polypeptide expressed by the aroE gene contained in NCIMB Deposit No. 40794;
(b) a polynucleotide complementary to the polynucleotide of(a); and
(c) a polynucleotide comprising at least 15 bases of the polynucleotide of (a) or (b).

8. A vector comprising the DNA of Claim 2.

9. A host cell comprising the vector of Claim 8.

10. A process for producing a polypeptide comprising: expressing from the host cell of Claim 9 a polypeptide encoded by said DNA.

11. A process for producing a cell which expresses a polypeptide comprising transforming or transfecting the cell with the vector of Claim 8 such that the cell expresses the polypeptide encoded by the cDNA contained in the vector.

12. A process for producing a aroE polypeptide or fragment comprising culturing a host of claim 9 under conditions sufficient for the production of said polypeptide or fragment.

13. A polypeptide comprising an amino acid sequence which is at least 70% identical to amino acid 1 to 257 of SEQ ID NO:2.

14. A polypeptide comprising an amino acid sequence as set forth in SEQ ID NO:2.

15. An antibody against the polypeptide of claim 13.

16. An antagonist which inhibits the activity of the polypeptide of claim 13.

17. A method for the treatment of an individual having need of aroE comprising: administering to the individual a therapeutically effective amount of the polypeptide of claim 13.

18. The method of Claim 16 wherein said therapeutically effective amount of the polypeptide is administered by providing to the individual DNA encoding said polypeptide and expressing said polypeptide *in vivo*.

19. A method for the treatment of an individual having need to inhibit aroE polypeptide comprising: administering to the individual a therapeutically effective amount of the antagonist of Claim 16.

20. A process for diagnosing a disease related to expression of the polypeptide of claim 13 comprising:
determining a nucleic acid sequence encoding said polypeptide.

21. A diagnostic process comprising:
analyzing for the presence of the polypeptide of claim 13 in a sample derived from a host.

22. A method for identifying compounds which bind to and inhibit an activity of the polypeptide of claim 13 comprising:

contacting a cell expressing on the surface thereof a binding for the polypeptide. said binding being associated with a second component capable of providing a detectable signal in response to the binding of compound to said binding. with a compound to be screened under conditions to permit binding to the binding; and
determining whether the compound binds to and activates or inhibits the binding by detecting the presence or absence of a signal generated from the interaction of the compound with the binding.

23. A method for inducing an immunological response in a mammal which comprises inoculating the mammal with aroE. or a fragment or variant thereof, adequate to produce antibody to protect said animal from disease.

24. A method of inducing immunological response in a mammal which comprises. through gene therapy. delivering

FIGURE 1. [SEQ ID NO:1]

```
   1  ATGAAGCTTG ATGGCTATAC ACGTTTAGCT GCAGTTGTTG CCAATCCTAT

  51  TAAGCATTCT ATTTCTCCCT TCATCCACAA TAGAGCCTTT GAGGCGACAG

 101  CTACCAACGG TGCTTATGTG GCTTGGGAGA TTGAAGCGAG TGACTTGGTA

 151  GAAACAGTGG CCAATATTCG TCGCTACCAG ATGTTTGGCA TCAATCTGTC

 201  CATGCCCTAT AAGGAGCAGG TGATTCCTTA TTTGGATAAG  CTGAGCGATG

 251  AAGCGCGCTT GATTGGTGCG GTTAATACGG TTGTCAATGA GAATGGCAAT

 301  TTAATTGGAT ATAATACAGA TGGCAAGGGA TTTTTTAAGT GCTTGCCTTC

 351  TTTTACAATT TCAGGTAAAA AGATGACCCT GCTGGGTGCA GGTGGTGCGG

 401  CTAAATCAAT CTTGGCACAG GCTATTTTGG ATGGCGTCAG TCAGATTTCG

 451  GTCTTTGTTC GTTCCGTTTC TATGGAAAAA ACAAGACCTT ACCTAGACAA

 501  GTTACAGGAG CAGACAGGCT TTAAAGTGGA TTTGTGTGCT TTAGAATATG

 551  TTTCTGAACT GCAAGCAAGG ATTGCCGAGT CGGATTTGCT AGTTAATGCC

 601  ACCAGTGTGG GCATGGATGG CCAATTCTCT CCTGTTCCTG AAAACATAGT
```

FIGURE 2. [SEQ ID NO:2]

```
   1  MKILDGYTRLA AVVANPIKHS ISPFIHNRAF EATATNGAYV AWEIEASDLV

  51  ETVANIRRYQ MFGINLSMPY KEQVIPYLDK LSDEARLIGA VNTVVNENGN

 101  LIGYNTDGKG FFKCLPSFTI SGKKMTLLGA GGAAKSILAQ AILDGVSQIS

 151  VFVRSVSMEK TRPYILDKLQE QTGFKVDLCA LEYVSELQAR IAESDLLVNA

 201  TSVGMDGQFS PVPENIVLPE TLLVADIIYQ PFETPFLKWA RSQGNPAVNG

 251  LGMLLYQ
```

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 0 854 188 A3

EUROPEAN PATENT APPLICATION

(12)

(88) Date of publication A3:
09.02.2000 Bulletin 2000/06

(43) Date of publication A2:
22.07.1998 Bulletin 1998/30

(21) Application number: **98300361.7**

(22) Date of filing: **20.01.1998**

(51) Int Cl.[7]: **C12N 15/53**, C12Q 1/68, C12N 15/62, C07K 16/12, G01N 33/566, A61K 38/44

(84) Designated Contracting States:
AT BE CH DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE

(30) Priority: **21.01.1997 US 36125 P**

(71) Applicant: **Smithkline Beecham**
Philadelphia, PA 19103 (US)

(72) Inventor: **Burnham, Martin K.R.,**
SmithKline Beecham Pharm.
Collegeville, PA 19426-0989 (US)

(74) Representative: **Mallalieu, Catherine Louise**
D. Young & Co.,
21 New Fetter Lane
London EC4A 1DA (GB)

(54) Streptococcus pneumoniae aroE polypeptides and polynucleotides

(57) The invention provides Streptococcus pneumoniae aroE polypeptides and DNA (RNA) encoding aroE polypetides and methods for producing such polypeptides by recombinant techniques. Also provided are methods for utilizing aroE polypeptide for the protection against infection, particularly bacterial infections.

EP 0 854 188 A3

Printed by Jouve, 75001 PARIS (FR)

European Patent Office

INCOMPLETE SEARCH
SHEET C

Application Number
EP 98 30 0361

Although claims 17-19, 23 and 24 are directed to a method of treatment of the human/animal body (Article 52(4) EPC), the search has been carried out and based on the alleged effects of the compound/composition.

Claim(s) not searched:
16,19

Reason for the limitation of the search:

Claims 16 and in part 19 relate to an antagonist of the polypeptide without giving a true technical characterization. Moreover, no such compounds are defined in the application. In consequence, the scope of said claims is ambiguous and vague, and their subject-matter is not sufficiently disclosed and supported (Art. 83 and 84 EPC). No search can be carried out for such purely speculative claims whose wording is, in fact, a mere recitation of the results to be achieved.

ANNEX TO THE EUROPEAN SEARCH REPORT ON EUROPEAN PATENT APPLICATION NO. EP 98 30 0361

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report. The members are as contained in the European Patent Office EDP file on The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

25-11-1999

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 5187071 A	16-02-1993	NONE	
WO 9533843 A	14-12-1995	US 5776736 A	07-07-1998
		EP 0763127 A	19-03-1997
WO 9818931 A	07-05-1998	AU 5194598 A	22-05-1998
		AU 6909098 A	22-05-1998
		EP 0942983 A	22-09-1999
		EP 0941335 A	15-09-1999
		WO 9818930 A	07-05-1998

EPO FORM P0459

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82